Finding planets and life among the stars

For the first time in history, mankind has the tools to look for other planets and whether these might host life

Margaret C Turnbull

Are we alone in the universe? How can we find out? What would happen if we made contact with extraterrestrial life forms, intelligent or otherwise? These questions have fascinated humans ever since we understood that there is a universe beyond our pale blue dot. As the British physicist and author Arthur C Clarke put it: “Two possibilities exist: either we are alone in the universe or we are not. Both are equally frightening”.

“Two possibilities exist: either we are alone in the universe or we are not. Both are equally frightening”

Speculation about the existence and nature of extraterrestrial life—and how humans might find and interact with it—has gone on for several millennia, but we live in the first age where the tools are available to collect, analyze and understand the potential evidence. The fundamental question “Are we alone?” drives the field of astrobiology, whose researchers tackle everything from microbes that thrive in extreme environments, to the origin of Earth’s moon and oceans, to the potential for ecosystems under the icy surfaces of Mars and Jupiter’s moon Europa. Momentum is also gathering for the search for habitable Earth-like worlds orbiting other stars. However, the detection of “exo-Earths”, and the discovery of any life signs thereupon, is one of the most technically and scientifically challenging feats ever attempted by astronomers.

“...the detection of “exo-Earths”, and the discovery of any life signs thereupon, is one of the most technically and scientifically challenging feats ever attempted by astronomers.”

To an astrobiologist, this image is also an illustration of life’s broadcast to the universe: “We are here”. Science fiction authors like to remind us that radio signals have now carried all manner of impressions about the human condition to anyone listening within a 100 light year radius—often exploring the potentially interesting reactions from our galactic neighbors. This sphere of human noise has now encountered perhaps five thousand planetary systems, about which we know very little.

What they sometimes forget, however, is that life on Earth has been broadcasting its presence to the stars in the form of reflected sunlight for far longer than a century. The light from our sun, reflected by Earth, is imprinted with the ever-changing colors of plants, oceans, oxygen, weather patterns, seasonal fluctuations, volcanoes, ice ages, tropical periods, and the rise and fall of biosystems over billions of years. Those signals, though faint, have now traveled far beyond the boundaries of our own Milky Way galaxy, past our sister Andromeda galaxy, beyond the Local Group of galaxies, past the boundaries of the Virgo supercluster of galaxies, and into the depths of a significant fraction of the observable universe. “We are here, and we are alive”, our planet has told the universe. Presumably, other planets have done the same, and in our quest for exo-Earths, that is the signal we seek.

Where do we look for other worlds that can support life as we know it? The universe is a very cold and largely inhospitable place, with the ambient temperature of space only a few degrees above absolute zero (−273.15 °Celsius). When new generations of stars form from giant galactic clouds of gas and dust, they become tiny nuclear heaters that keep their local environments warm enough
for nearby planets to have liquid water and enough energy to drive the complex chemical interactions that make life possible.

Our sun is a typical dwarf star, moving through an ocean of several hundred billion other stars in the Milky Way galaxy. Around each star, there is what is called a “habitable zone”, where a planet like the Earth is at just the right temperature for liquid water to exist on its surface (Fig 2). That zone is huddled relatively tightly around the sun, stretching from just about Earth’s orbit, to somewhere beyond the orbit of Mars. The very young Venus may have had water oceans much like the Earth, but they quickly evaporated and were lost to space owing to the intense radiation from the sun at that distance. Meanwhile, Mars, which shows clear signs of ancient rivers and seas, is simply too small to retain a substantial atmosphere, and whatever water was not frozen onto the surface was eventually lost to space too. In theory, if the young Venus and Mars could have traded places, we might have had two habitable worlds in our Solar System today.

We know that the sun is a star, and all stars are suns—but every star is unique. Our galaxy is dominated by stars much smaller and fainter than the sun. These so-called red dwarfs are so faint that none can be seen with the naked eye, even in the darkest of skies. The vast majority of the universe’s habitable zones surround these tiny invisible stars, which burn through their meager fuel so slowly that their habitable zones might remain stable for trillions of years.

Red dwarfs have long been interesting to astrobiologists in terms of their suitability for supporting biological life for two main reasons. First, their habitable zones are so close to the star (1/4 the orbit of Mercury) that a planet in orbit there would almost certainly be tidally locked. This means that the same side of the planet always faces the star, such that one side of the planet will be a blazing noonday Sahara, while the other is a freezing midnight Siberia. Scientists originally feared that such a situation would be unstable and would tend to condense all of the planet’s moisture, indeed the entire planetary atmosphere, into a giant icecap on the night side. Second, red dwarfs are known for intense flaring activity, which would potentially sterilize the planet’s dayside surface (Fig 3). These two factors might spell doom for Earth-sized worlds around red dwarfs, but if slightly larger planets called Super Earths are common, their thicker and more protective atmospheres could solve both problems at once and rescue billions of stars from seeming infertility.

Meanwhile, stars more massive than the sun are hotter, bluer, and exponentially more luminous. These stars, although relatively few in number, are so bright that they are the source of most of our galaxy’s light. The habitable zones of these stars are farther away from the star and wider—wide enough for several habitable planets to orbit therein. However, stars even just a few times the mass of the sun are surprisingly short-lived. Although they have more hydrogen to burn, but with all of that mass pressing down on their interiors causes their fusion engines to run hot enough to consume all of that fuel, even before the planets around them finish forming. Whereas a star like the sun will live for ten billion years, more massive stars may not survive for even one billion years. At that point, these stars become unstable, shedding mass in all directions as they swell into giants, engulfing and destroying whatever young planets might be forming there.

The different behavior of differently sized stars gives rise to a lovely relationship between massive stars, their smaller sun-like counterparts, and life. Before the first stars formed, our universe contained nothing in the way of elements like iron, nickel, oxygen, carbon, and nitrogen—the things of which we and our planet are made. When the first massive stars formed, their internal nuclear furnaces fused copious amounts of hydrogen into helium, and when they ran out of hydrogen, they fused helium into oxygen, carbon, and nitrogen, and when they ran...
out of helium, they burned ever-heavier elements, making nickel and silicon, all the way up to iron. Then, they exploded, and in the process of exploding, they gave rise to every other element found in the universe today. These massive stars are the manufacturers and distributors that supply our entire galaxy with the elements required for planets and life. Subsequent generations of lower mass stars could then support living ecosystems for billions of years, some of which evolved into technological civilizations like ours.

Moving beyond the edge of our galaxy, we find similar oases, each of them twirling with a hundred billion stars and becoming increasingly enriched over time with molecules essential for life. NASA’s Hubble Space Telescope has discovered that every tiny patch of sky, about the size of the eye of a needle held at arm’s length, is actually filled with thousands of galaxies suspended between us and the edge of the observable universe. This appears to be the case no matter which direction one looks, implying that there are as many galaxies in the observable universe as there are stars in our own galaxy. In other words, if stars are the sites for life, then there are at least a hundred billion potential places for worlds like ours (Fig 4).

When it comes to seeing ourselves in the cosmos, the current generation of young adults is different from their ancestors. They have grown up in a world where planetary systems outside our own are discovered and announced on an almost weekly basis. Since 1995, humanity has gone from zero examples of extrasolar worlds to more than five thousand—almost 500 of which reside in multi-planet systems.

“...if stars are the sites for life, then there are at least a hundred billion potential places for worlds like ours”

The first planets discovered were all larger than Jupiter, orbiting very close to their suns, for the simple reason that these planets were the easiest to find with state of the art instrumentation. Astronomers have used the so-called radial velocity method to detect the periodic red- and blue-shifting of starlight as the star and planet orbit around their common center of mass. This effect is larger for more massive planets and easiest to confirm when the period is very short. In this context, “easiest” should be taken to mean many years of painstaking, isolating work that was greeted with strong skepticism and even mockery by the wider astronomical community.

With the radial velocity method, the planet itself remains invisible, but its effect on the star is what we see; the detection is indirect. The only information we can gather about a planet from this method is an estimate of its mass and how far away from the star it sits. The latter helps us to guess at the planet’s temperature, but we cannot know the physical size of the planet, its density, composition, structure, or anything definite about its climate and the potential for life without some additional information. Nevertheless, hundreds of planets have been detected in this way, eventually allowing astronomers to locate smaller and smaller worlds at farther and farther distances from their stars. To date, one of the smallest planets found using the radial velocity method is practically sitting in our own backyard. The nearest star system to the Sun, Alpha Centauri, consists of three stars, one of which is now believed to harbor a planet slightly larger—but much hotter—than the Earth (Fig 5; [1]).

The hard-won success of radial velocity searches, and the sheer diversity of planetary systems that it revealed, gave impetus to NASA’s Kepler mission—the first space telescope dedicated solely to detecting planets more like ours. Kepler uses a different detection technique, which only works for planets whose orbits happen to be edge-on, as seen from the Earth. As the planet passes in front of its star, the star appears to become slightly fainter, returning to its original brightness once the planet has moved on. This gentle wink, called a transit, lasts a few hours and can be seen across the galaxy. Transits reveal the size of a planet and how far from the star it sits. Again, the latter helps us to guess at the planet’s temperature; we still cannot know the mass of the planet, its density, composition, structure, or anything definite about its climate and the potential for life. Nevertheless, Kepler is indisputably the most successful planet hunting effort ever executed, with thousands of discoveries to its credit.

Kepler’s harvest of alien worlds has naturally been subject to much speculation regarding their habitability. Both the radial
velocity and transit techniques select for warm planets near their stars. By combining Kepler data with radial velocity measurements, however, astronomers have now identified a few handfuls of planets that seem to be about the same density as the Earth—implying a solid rocky surface much like that beneath our feet—while others seem more likely to be blanketed with a thick atmosphere or deep ocean. Intriguingly, some of these planets also appear to be in their stars’ habitable zones. At the time of writing, the Planetary Habitability Laboratory at the University of Puerto Rico highlights 23 planets that they believe have a chance of supporting life as we know it (Fig 6; http://phl.upr.edu).

However, assessing a planet for habitability and life is a difficult business when all you have is its size and an approximate temperature. For example, we know that the Earth’s oceans were formed from impacts with icy asteroids, stirred up by the giant planets farther out in the Solar System. Did these other planets have a water delivery service, too, or are they dry? We also know that plate tectonics on Earth are crucial to cycling the atmosphere and keeping our climate stable over billions of years. The occurrence of plate tectonics on the Earth may well be related to a huge impact long ago that resulted in the formation of our large Moon, leaving the Earth’s crust a bit thinner in the process [2]. Do these other worlds have a similar way of cycling carbon between the atmosphere and rocks?

These planets also have other characteristics that we do know about and that only adds to the habitability mystery. For example, nearly all of them are so-called Super Earths, up to three times the size of our own planet [3]. Having no analog for such a planet in the Solar System, we wonder what their atmospheres and weather patterns are like? Are they just extra large versions of the Earth, or extra small versions of Neptune? Do they have solid surfaces? Are they shrouded in dense cloud cover, or can sunlight make it to the surface, where photosynthetic organisms might harvest it?

Furthermore, because Kepler preferentially finds planets in tight orbits, most of these habitable zone planets are orbiting red dwarfs. Kepler 186-f, the first planet found to be nearly the same size of the Earth and in the habitable zone, is one such example. As described above, we expect these planets to be tidally locked. What does that mean for life? On Earth, much of biological evolution has been driven by the need to move and adapt to the periodic changes of day, night, and seasons. If one is immersed either in eternal day, or endless night, where nothing ever changes, is there still a motivation to move? To develop fins, legs, wings, and civilizations?

The above landslide of data, while raising many questions, tells us two important things: planetary systems are very common among the sun’s neighbors, and many of them look nothing like the planets in our Solar System. So how can we move beyond beautiful but speculative artists’ renditions and begin to discern what these planets are truly like? To achieve the “holy grail” of detecting and understanding the nearest Earth-sized planets in terms of habitability and life, we need to take the next step: direct imaging.

Direct imaging is easy to explain, because it involves nothing more than simply taking a picture of a planetary system around a nearby star. In 1990, on its way out of the Solar System, the Voyager 1 probe turned around and took a picture of the Earth from roughly 10 billion kilometers away—about 1/1,500th of the way to the nearest star. From that distance, the Earth appears not as a marble, but as a speck of light where all spatial information about continents,
oceans, and life is blended together into a single pixel. Carl Sagan described Voyager’s image of the Earth as a “pale blue dot”, a very different impression from the detailed tapestry of the Blue Marble (Fig 7).

And yet, that tiny pinprick of reflected sunlight is laden with valuable information. We think of the Earth’s atmosphere as being transparent, but specific wavelengths of the sun’s light are actually scattered and absorbed in different ways by the different components of the atmosphere. Nitrogen molecules preferentially scatter short wavelengths, giving us a blue sky. Water vapor and carbon dioxide molecules soak up so much light in the infrared that astronomers have difficulty seeing the stars at certain wavelengths. Methane, generated primarily by bacteria, also has a unique signature in the infrared. Meanwhile, oxygen has a sharp, distinctive feature in the green, and our protective ozone layer turns the entire planet black in the ultraviolet. The simultaneous presence of abundant oxygen and methane is especially interesting, because those two gases should quickly react and disappear into carbon dioxide and water. Methane could be continually replenished by volcanic activity, but there is no clear way to sustain large amounts of oxygen in the air—except by life itself, in the form of photosynthetic algae and plants (Fig 8; [4]). Our atmosphere is like a signaling mirror, indicating the presence of life.

“Taking a picture from another planet is by far the most difficult problem ever tackled by observational astronomy”

Earth’s surface is also sending signals. As the world turns, it presents its various faces: oceans, forests, grasslands, glaciers, and clouds. All of these surfaces reflect different amounts of sunlight at different wavelengths, such that the color of that pale dot is not always the same shade of blue from hour to hour, week to week, or season to season. Plants in particular have a distinctive signature, being very dark at visible wavelengths, and suddenly brightening in the infrared. By monitoring the changing spread of colors emanating from our tiny dot over time, an astronomer looking back at us from a nearby star could in principle tell that this planet has a 24-hour day, weather patterns, seasons, water, active geology, and life. In this context, a picture is worth much more than a thousand words.

If so much insight about distant planets is waiting for us in a simple picture, then what is stopping us from taking one? The answer is that separating the tiny light of a planet from the overpowering blaze of its star is remarkably hard. Habitable planets are nestled very close to their stars, and they are ten billion times fainter than the ball of fire they orbit. Some planetary systems, like our own, are also filled with dust from collisions between asteroids. As Aki Roberge from NASA’s Goddard Space Flight Center recently put it, “Observing an exo-Earth is like seeing a dinner candle, a hair’s width away from the Luxor Sky Beam, from two football fields away, in a dust storm”. Worse, that dinner candle is flickering in a sea of background candles, and it is not always easy to tell which is which.

Taking a picture from another planet is by far the most difficult problem ever tackled by observational astronomy, akin to the most extreme high-contrast photography ever attempted. Such a feat can only be achieved from space, where there is no terrestrial atmosphere to interfere. In 2013, NASA commissioned two teams of scientists and engineers to find a solution to these challenges, and to design a planet-imaging mission with a price tag of less than US$1 billion (http://exep.jpl.nasa.gov/std4/). Each team is currently exploring a different way of achieving a common objective: find a way to block the light of nearby stars, while still allowing the light of their planets to be detected.

The problem is that light propagates through the universe as a wave. Although telescopes are designed to focus these waves into crisp images on a detector, waves have the unfortunate habit of diffracting around every surface they pass through.
encounter, such that starlight ends up spread all over the resulting image. One way of controlling this is to alter the size and/or shape of the telescope, such that these waves of starlight are either concentrated into ultra fine points of light, or are herded completely out of the picture, leaving the light from planets behind. This is the tactic of the first NASA team. Their design, called a “corona-graph”, utilizes a specially designed telescope housing a train of mirrors and lenses that, if all goes well, could allow for the detection of habitable zone planets around several of the nearest, brightest stars.

The second NASA team is trying a different approach: block the starlight before it even hits the telescope. This scenario requires two spacecrafts. The first carries a high-quality but otherwise run of the mill telescope, with no special optics required. The second spacecraft is a gigantic opaque screen, known as a “starshade”. This simple concept is familiar to anyone who has shielded their eyes while driving westward into the evening sun, in order to see other vehicles, bicycles, and people. The starshade, when aligned between the telescope and a bright target star, casts a deep shadow.
over the telescope, allowing the faint signal from any orbiting exoplanets to emerge (Fig. 9).

Naively, one might imagine that a simple round disk would do the job, as long as it is large enough to cast a shadow that the telescope can float within. But it turns out that light waves encountering a round disk will diffract around it in such a way that they merge on the other side of it to create a bright central spot, almost as if the starshade were not even there. This behavior is the primary reason that starshades have not been used since the idea was first proposed by Lymann Spitzer of Princeton University 50 years ago. But in the last decade, starshade engineers have identified a family of precise mathematical shapes, looking much like a sunflower with twenty or thirty petals that cause starlight to diffract in a more favorable pattern. This type of starshade is capable of creating a shadow so dark that only one in ten billion stellar photons makes it through—just enough to find habitable worlds orbiting forty or fifty of the brightest sun-like stars in the sky.

Each of these space observatory designs has advantages and disadvantages. The coronagraph is a more familiar architecture, having heritage with NASA’s Hubble Space Telescope and with ESA’s sun-observing SOHO mission. Meanwhile, the starshade idea is only beginning to gain acceptance as a realistic option in astronomical circles. But the large number of specialized optics required for the coronagraph means that, for a price tag of just under US$1 billion, its planet-finding performance is unlikely to match that of a starshade. On the other hand, the starshade must be flown all around the sky, taking days to weeks to line up with stars in different directions, while avoiding dangerous angles that reflect blinding sunlight back into the telescope. This makes it hard to take a second look at systems where something interesting was found, necessitating onboard instrumentation that can distinguish planets from other sources and identify signs of habitability, in a single image.

In the end, perhaps NASA will find a way to hybridize these concepts and fly an exoplanet imaging mission that is more cost effective, high performance, efficient, and flexible than either concept alone. To be compelling, such a program must provide for two things: pure discovery—finding planets previously unheard of—and for learning much more about the planets than can be gleaned from radial velocity and transit surveys.

If one thing can be said about planets, it is that they are vastly more complicated than their parent stars. From transit and radial velocity surveys, we already know that they span every possible size and temperature, and we know that multi-planet systems can have very different arrangements from our Solar System. From geology, we also know that the Earth’s current appearance has changed dramatically over time, such that the Earth does not look like it did a billion years ago. Our world was shaped by many unpredictable events: the formation of a large moon, the scattering of icy bodies into the inner solar system, and extensive alterations by biological activity. From planetary science, we know that the planets, moon, and minor bodies throughout our Solar System had their current locations and structure entirely determined by unpredictable events, including our global biosystem, which is both a direct cause and consequence of everything that happened here on Earth, including life.

But does that mean that there are no other habitable or inhabited worlds among the sun’s nearest neighbors? Are the events that have given rise to life on Earth really so rare? Biologists have found that there is essentially no corner of this globe, however, environmentally harsh, that is not home to
some living thing. Microbial life is so tenacious that it is notoriously difficult to maintain a sterile environment, even when subjecting creatures to radiation, starvation, and desiccation. Life seems to have taken hold of this planet when it was still a newborn, and it has survived and thrived in spite of cataclysmic events over billions of years. Knowing this, perhaps the presence of life on other worlds should not be surprising at all, even if there is no other world quite like ours.

Every aspect of the search for alien life reminds us of the Earth’s smallness. The incomprehensible number of stars, the vast distances between them, the utter emptiness and bitter cold of interstellar space, the billion year timescales over which planets are born and evolve—all of this seem to indicate our own insignificance in the grand scheme of things. But consider for one moment that there are more living organisms on Earth than there are stars in the entire universe. This unimaginable wealth of life swims in the deepest oceans, roams across continents, flies over the highest mountains, and tenaciously inhabits the most hostile environments. Can we really ask for more than that?

And yet, there is more. An event of cosmic significance is happening on this little dot: One species is coming into a difficult, sometimes heartbreaking, but also promising consciousness of its own ability to either maintain or destroy the world it depends upon. As a result, that species must also grapple emotionally with the paradox of being simultaneously both very small and very significant. Soon, that species may realize that it inhabits a planet that is both entirely unique, and yet just one of a billion living worlds.

“Life seems to have taken hold of this planet when it was still a newborn, and it has survived and thrived in spite of cataclysmic events over billions of years.”

If one tiny speck in the universe can be filled with such richness, expressing so much diversity of beauty, struggle, and adaptation, what will we discover on other worlds? As Carl Sagan once reminded us, we are a nomadic species, and our ancestors crossed continents and oceans in search of new opportunities. Our descendants may well cross interstellar space, diving deep into the arms of the Milky Way. Will they one day “gaze up and strain to find the pale blue dot in their skies”, marveling “how perilous our infancy, how humble our beginnings”, as Sagan imagined? The quest to find other worlds, and the effort to understand them, will transform us into the kind of species that is both capable and worthy of calling this magnificent living universe our home.

Conflict of interest
The author declares that she has no conflict of interest.

References